

DOI: 10.3969/j.issn.1005-8982.2023.19.010
文章编号: 1005-8982 (2023) 19-0066-08

临床研究·论著

多因素 Logistic 回归模型和 XGBoost 模型对舌癌患者放疗期间发生口腔感染的预测价值*

张洋¹, 赵化荣¹, 刘攀¹, 宿伟鹏¹, 龚忠诚²

(新疆医科大学第一附属医院 1. 肿瘤中心, 2. 颌面肿瘤外科, 新疆 乌鲁木齐 830054)

摘要: **目的** 探讨多因素 Logistic 回归模型和 XGBoost 模型对舌癌患者放疗期间发生口腔感染的预测价值。**方法** 选取2003年1月—2022年12月在新疆医科大学第一附属医院就诊的431例放疗的舌癌患者, 随机分为训练组288例和预测组143例。比较多因素 Logistic 回归模型和 XGBoost 模型对舌癌患者放疗期间发生口腔感染的预测效能。**结果** 多因素 Logistic 回归模型结果显示, 年龄 $[\hat{OR} = 3.250 (95\% \text{ CI}: 1.476, 7.634)]$, 肿瘤分期 $[\hat{OR} = 2.941 (95\% \text{ CI}: 1.248, 7.613)]$, 口腔环境 $[\hat{OR} = 0.210 (95\% \text{ CI}: 0.079, 0.502)]$, 是否手术 $[\hat{OR} = 0.285 (95\% \text{ CI}: 0.113, 0.663)]$, 血红蛋白 $[\hat{OR} = 0.323 (95\% \text{ CI}: 0.139, 0.712)]$, 血清白蛋白 $[\hat{OR} = 0.353 (95\% \text{ CI}: 0.148, 0.851)]$ 是放疗期间发生口腔感染的独立预测因素。XGBoost 模型结果显示, 口腔环境、手术、肿瘤分期、血清白蛋白、年龄、同步化疗、红细胞计数、血红蛋白、中性粒细胞计数为重要性指标。多因素 Logistic 回归模型和 XGBoost 模型受试者工作特征曲线下面积分别为0.830、0.835, 两者比较, 差异无统计学意义($P > 0.05$); 敏感性分别为88.24%(95% CI: 0.729, 1.000)、82.35%(95% CI: 0.642, 1.000); 特异性分别为68.25%(95% CI: 0.601, 0.764)、69.84%(95% CI: 0.627, 0.786)。**结论** 多因素 Logistic 回归模型和 XGBoost 模型对舌癌患者放疗期间发生口腔感染的预测均有意义, 两者预测效能相当。建立模型有助于筛选出口腔感染的高危人群, 及早采取预防措施, 降低口腔感染发生风险。

关键词: 舌癌; 放疗; 口腔感染; 多因素 Logistic 回归模型; XGBoost 模型

中图分类号: R739.86

文献标识码: A

Predictive value of logistic regression model and XGBoost model for the occurrence of oral infection during radiotherapy in patients with tongue cancer*

Zhang Yang¹, Zhao Hua-rong¹, Liu Pan¹, Su Wei-peng¹, Gong Zhong-cheng²

(1. Oncology Center, 2. Oncology Department of Oral & Maxillofacial Surgery, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang 830054, China)

Abstract: Objective This study aimed to explore the predictive value of multifactor logistic regression and XGBoost models for oral infections during radiotherapy in tongue cancer patients. **Methods** A total of 431 tongue cancer patients receiving radiotherapy at the First Affiliated Hospital of Xinjiang Medical University from January 2003 to December 2022 were randomly divided into a training group ($n = 288$) and a prediction group ($n = 143$). The predictive performance of multifactor logistic regression and XGBoost models for oral infections during radiotherapy in tongue cancer patients was compared. **Results** The results of the multifactor logistic regression model showed that age $[\hat{OR} = 3.250 (95\% \text{ CI}: 1.476, 7.634)]$, tumor stage $[\hat{OR} = 2.941 (95\% \text{ CI}: 1.248, 7.613)]$,

收稿日期: 2023-04-11

* 基金项目: 省部共建国家重点实验室(No.: SKL-HIDCA 2021-47, SKL-HIDCA 2021-6, SKL-HIDCA 2022-44)

[通信作者] 龚忠诚, E-mail: gump0904@aliyun.com; Tel: 0991-4366081

oral environment [$\hat{OR} = 0.210$ (95% CI: 0.079, 0.502)], surgery status [$\hat{OR} = 0.285$ (95% CI: 0.113, 0.663)], hemoglobin [$\hat{OR} = 0.323$ (95% CI: 0.139, 0.712)], and serum albumin [$\hat{OR} = 0.353$ (95% CI: 0.148, 0.851)] were independent predictors of oral infections during radiotherapy. The XGBoost model identified oral environment, surgery status, tumor stage, serum albumin, age, concurrent chemotherapy, red blood cell count, hemoglobin, and neutrophil count as important predictors. The area under the receiver operating characteristic curve (AUC) for the multifactor logistic regression model and the XGBoost model were 0.830 and 0.835, respectively, with no statistically significant difference between them ($P > 0.05$). Sensitivity was 88.24% (95% CI: 0.729, 1.000) and 82.35% (95% CI: 0.642, 1.000), while specificity was 68.25% (95% CI: 0.601, 0.764) and 69.84% (95% CI: 0.627, 0.786) for the multifactor logistic regression and XGBoost models, respectively. **Conclusion** Both multifactor logistic regression and XGBoost models have significant predictive value for oral infections during radiotherapy in tongue cancer patients, with comparable predictive performance. Establishing these models can help identify high-risk individuals for oral infections, enabling early preventive measures and reducing the risk of oral infections.

Keywords: tongue neoplasms; radiotherapy; oral infection; logistic regression model; XGBoost model

舌癌是头颈部恶性肿瘤中最常见的一种,发病率逐年上升,手术和放化疗是舌癌的主要治疗手段^[1-2]。放疗可降低舌癌患者局部复发风险,但放疗期间容易发生口腔感染,发生率为 19.3%~41.0%。口腔感染可造成患者进食疼痛、体重下降、降低生活质量,严重者导致放疗中断,影响疗效,并增加治疗费用,延长住院时间。如何能降低患者口腔感染风险是治疗过程中医护人员关注的问题之一^[3-6],大多数研究探讨口腔肿瘤与口腔感染间的危险因素分析,未能构建有效的预测模型^[7-9]。本文拟对舌癌患者放疗期间发生口腔感染的危险因素进行分析,应用 R 软件整合相关危险因素,建立最佳预测模型,精准识别舌癌患者放疗期间发生口腔感染的高风险人群,有利于医护人员预见性地采取预防措施。

1 资料与方法

1.1 一般资料

选取 2003 年 1 月—2022 年 12 月在新疆医科大学第一附属医院接受放疗的舌癌患者 431 例。纳入标准:①病理明确诊断为舌恶性肿瘤;②舌癌或舌癌术区第一次接受放疗;③临床资料完整,放疗前无口腔感染。排除标准:①合并第 2 原发肿瘤,合并严重的脏器功能不全。②放疗前 1 个月内服用糖皮质激素、免疫抑制剂及抗菌药物等影响口腔感染的药物。将入组患者随机分为训练组和预测组,训练组用于构建预测模型,预测组用于模型效果评价。其中训练组 288 例,女性 138 例,男性 150 例;年龄 ≥ 60 岁 147 例,年龄 < 60 岁 141 例;I、II 期 124 例,III、IV 期 164 例;手术 111 例,非手术 177 例。预测组 143 例,

女性 72 例,男性 71 例;年龄 ≥ 60 岁 64 例,年龄 < 60 岁 79 例;I、II 期 70 例,III、IV 期 73 例;手术 59 例,非手术 84 例。

1.2 治疗方法

①放疗:取仰卧位,颈肩架固定,使用 CT 模拟机扫描定位,根据 CT 扫描图像勾画靶区和危及器官。采用瓦里安直线加速器 X 射线照射进行适形或调强放射治疗。②根治性放疗剂量:pGTVp(影像所见大体肿瘤)及 pGTVn(转移淋巴结)总剂量 66 Gy/2.0 Gy/33 f,PTV1(高危临床靶区)总剂量 60.06 Gy/1.82 Gy/33 f,PTV2(低危临床靶区)总剂量 54.12 Gy/1.64 Gy/33 f,1 次/d,放疗 5 次/周。③术后辅助放疗剂量:PTV1(高危临床靶区)总剂量 60 Gy/2.0 Gy/30,PTV2(低危临床靶区)总剂量 54 Gy/1.8 Gy/30 f,1 次/d,放疗 5 次/周。

1.3 诊断标准

当患者放疗期间出现口腔炎症时,取口腔分泌物,完善细菌或真菌涂片,行细菌培养、真菌培养及药敏实验,根据病原学检测结果和临床医师的经验判断患者是否发生口腔感染。根据 RTOG 急性放射性损伤分级标准,将放射性口腔黏膜炎分为 5 级:0 级,黏膜无变化;1 级,黏膜充血,可有轻度疼痛;2 级,出现片状黏膜炎或者伪膜形成,中度疼痛需止痛治疗;3 级,出现融合的黏膜炎或者伪膜形成,轻度诱发出血;4 级,出现溃疡,出血,坏死,或者危及生命。入组患者放射性口腔黏膜炎发生程度为 2~4 级。

1.4 研究方案

根据放疗期间口腔感染与否,将舌癌患者分为

口腔感染组与未感染组。对入组患者的指标进行统计学分析,包括:基本信息资料(性别构成、年龄、教育水平、体质指数、吸烟、饮酒、口腔环境);临床相关资料(肿瘤分期、同步化疗、是否手术,高血压、糖尿病、Kamofsky 评分);实验室相关指标(白细胞计数、中性粒细胞计数、血红蛋白、红细胞计数、血清白蛋白)。

1.5 统计学方法

数据分析采用 SPSS 和 R-Studio 软件(4.2.2 版)。计数资料以构成比或率(%)表示,比较用 χ^2 检验;多因素 Logistic 回归分析筛选独立危险因素,并构建 Logistic 预测模型;用 XGBoost 中“xgb.importance”对预测模型变量进行重要性排序,并构建 XGBoost (eXtreme Gradient Boosting)模型;绘制受试者工作特

征(receiver operating characteristic, ROC)曲线。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 舌癌患者放疗期间发生口腔感染的单因素分析

训练组 288 例舌癌患者放疗期间发生口腔感染 42 例,发生率为 14.58% (42/288),未出现感染 246 例。两组性别构成、教育水平、体质指数、高血压患病率、糖尿病患病率、吸烟率、饮酒率、Kamofsky 评分、白细胞计数比较,经 χ^2 检验,差异均无统计学意义($P > 0.05$);两组年龄、肿瘤分期、口腔环境、是否手术、同步化疗,中性粒细胞计数,血红蛋白,红细胞计数,血清白蛋白比较,差异均有统计学意义($P < 0.05$)。见表 1。

表 1 舌癌患者放疗期间感染组与非感染组不同因素间口腔感染率的比较

组别	n	男/女/例		年龄 例(%)		教育水平 例(%)			体质指数 例(%)		
				< 60 岁	≥ 60 岁	大专及以上学历	高中	高中以下	<18.5 kg/m ²	18.5~23.9 kg/m ²	>23.9 kg/m ²
感染组	42	20/22		12(28.57)	30(71.43)	13(30.95)	8(19.05)	21(50.0)	2(4.76)	25(59.52)	15(35.71)
非感染组	246	130/116		129(52.44)	117(47.56)	58(23.58)	51(20.73)	137(55.69)	13(5.28)	136(55.28)	97(39.43)
χ^2 值		0.211		7.251			1.053			0.263	
P 值		0.646		0.007			0.591			0.877	

组别	肿瘤分期 例(%)		高血压 例(%)		糖尿病 例(%)		吸烟 例(%)		饮酒 例(%)	
	I、II	III、IV	否	是	否	是	否	是	否	是
感染组	8(19.05)	34(80.95)	30(71.43)	12(28.57)	34(80.95)	8(19.05)	21(50.0)	21(50)	20(47.62)	22(52.38)
非感染组	116(47.15)	130(52.85)	190(77.24)	56(22.76)	214(86.99)	32(13.01)	141(57.32)	105(42.68)	127(51.63)	119(48.37)
χ^2 值	10.441		0.387		0.647		0.511		0.098	
P 值	0.001		0.534		0.421		0.475		0.754	

组别	口腔环境 例(%)		手术 例(%)		同步化疗 例(%)		Kamofsky 例(%)		白细胞计数 例(%)	
	良	差	否	是	否	是	<70 分	≥70 分	<4.0 × 10 ⁹ /L	≥4.0 × 10 ⁹ /L
感染组	8(19.05)	34(80.95)	33(78.57)	9(21.43)	27(64.29)	15(35.71)	25(59.52)	17(40.48)	7(16.67)	35(83.33)
非感染组	120(48.78)	126(51.22)	144(58.54)	102(41.46)	206(83.74)	40(16.26)	137(55.69)	109(44.31)	24(9.76)	222(90.24)
χ^2 值	11.669		5.263		7.574		0.087		1.784	
P 值	0.000		0.022		0.006		0.768		0.183	

组别	中性粒细胞计数 例(%)		红细胞计数 例(%)		血红蛋白 例(%)		血清白蛋白 例(%)	
	<1.5 × 10 ⁹ /L	≥1.5 × 10 ⁹ /L	<4.3 × 10 ¹² /L	≥4.3 × 10 ¹² /L	<130 g/L	≥130 g/L	<35 g/L	≥35 g/L
感染组	6(14.29)	36(85.71)	27(64.29)	15(35.71)	30(71.43)	12(28.57)	14(33.33)	28(66.67)
非感染组	7(2.85)	239(97.15)	87(35.37)	159(64.63)	106(43.09)	140(56.91)	27(10.98)	219(89.02)
χ^2 值	10.893		11.366		10.451		12.914	
P 值	0.005		0.000		0.001		0.000	

2.2 舌癌患者放疗期间发生口腔感染的多因素 Logistic 回归分析

将表 1 中单因素分析差异有统计学意义 ($P < 0.05$) 的 9 个变量纳入多因素 Logistic 回归分析中, 结果显示: 年龄 [$\hat{OR} = 3.250$ (95% CI: 1.476, 7.634)]、肿瘤分期 [$\hat{OR} = 2.941$ (95% CI: 1.248, 7.613)]、口腔环境 [$\hat{OR} = 0.210$ (95% CI: 0.079, 0.502)]、是否手术 [$\hat{OR} = 0.285$ (95% CI: 0.113, 0.663)]、血红蛋白 [$\hat{OR} = 0.323$ (95% CI: 0.139, 0.712)]、血清白蛋白 [$\hat{OR} = 0.353$, (95% CI: 0.148, 0.851)] 是舌癌患者放疗期间口腔感染的危险因素 (见表 2)。采用条件 Logistic 逐步回

归, 纳入标准 0.05, 排除标准 0.10。基于多因素 Logistic 回归分析差异有统计学意义的 6 项独立预测因素建立舌癌患者放疗期间口腔感染的预测模型。 $Z = -2.566 + 1.179 \times (\text{年龄}) + 1.079 \times (\text{肿瘤分期}) + 1.560 \times (\text{口腔环境}) - 1.254 \times (\text{手术}) - 1.129 \times (\text{血红蛋白}) - 1.041 \times (\text{血清白蛋白})$, 并绘制舌癌患者放疗期间发生口腔感染的列线图 (见图 1)。风险列线图上方附有比例尺, 每个因素对应比例尺得到各自得分, 所有预测因素单项评分之和为总分, 总分为 463 分, 每位患者各个因素相加总分越高, 提示放疗期间发生口腔感染风险的概率越大。

表 2 舌癌患者放疗期间口腔感染预测因素的多因素 Logistic 回归分析参数

自变量	b	S_b	Wald χ^2	P 值	\hat{OR}	95% CI	
						下限	上限
年龄	1.179	0.416	2.831	0.005	3.250	1.476	7.634
肿瘤分期	1.079	0.457	2.363	0.018	2.941	1.248	7.613
口腔环境	1.560	0.466	3.347	0.001	4.760	1.993	12.603
手术	-1.254	0.447	-2.806	0.005	0.285	0.113	0.663
血红蛋白	-1.129	0.413	-2.733	0.006	0.323	0.139	0.712
血清白蛋白	-1.041	0.443	-2.350	0.019	0.353	0.148	0.851

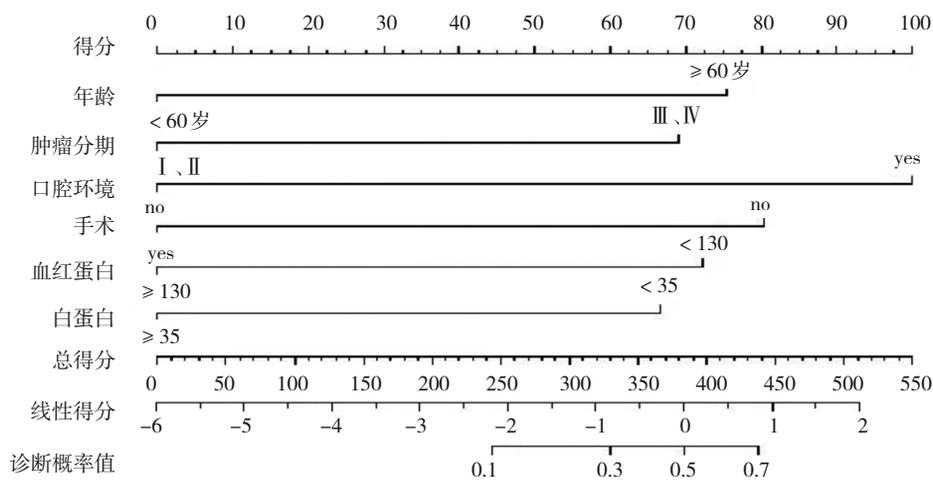


图 1 舌癌患者放疗期间发生口腔感染的风险预测列线图

2.3 建立预测舌癌患者放疗期间口腔感染的 XGBoost 模型

将训练集纳入 XGBoost 反复训练得到最佳参数 ($\text{max_depth} = 6, \text{eta} = 0.3, \text{nthread} = -1, \text{objective} = \text{binary: Logistic}, \text{nrounds} = 10$), 建立模型。用 XGBoost 算法, 最大对数似然指数 = 0.522 时, 模型敏感性为 82.35% (95% CI: 0.642, 0.100), 特异性为 69.84% (95% CI:

0.627, 0.786), ROC 曲线下面积 (area under the curve, AUC) 为 0.835 (95% CI: 0.741, 0.929)。见图 2。纳入 XGBoost 模型的所有指标中, 按重要程度排列, 依次排在前 9 位的指标如下图 (见图 2)。

2.4 2 个预测模型效能比较

采用测试集数据比较 2 个模型对舌癌患者放疗期间发生口腔感染风险的预测效果。训练组和预

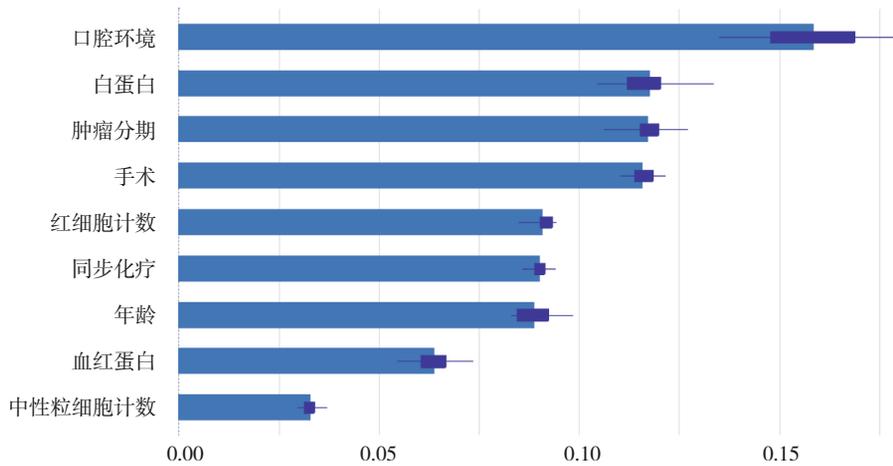


图 2 XGBoost模型中各变量重要性特征排列

测组数据集之间各项临床指标比较,差异均无统计学意义($P>0.05$)(见表 3)。XGBoost 预测模型的敏感性、阳性预测值、阴性预测值低于多因素 Logistic 回归模型,而准确性、特异性高于多因素 Logistic 回

表 3 训练组和预测组两组数据集之间各项临床指标比较

组别	n	男/女/ 例	感染 例(%)		教育水平 例(%)			体质量指数 例(%)		
			否	是	大专及以上	高中	高中以下	< 18.5 kg/m ²	18.5~23.9kg/m ²	>23.9kg/m ²
训练组	288	150/138	246(85.42)	42(14.58)	71(24.65)	59(20.49)	158(54.86)	15(5.21)	161(55.9)	112(38.89)
预测组	143	71/72	126(88.11)	17(11.89)	30(20.98)	29(20.28)	84(58.74)	10(6.99)	84(58.74)	49(34.27)
χ^2 值		0.139	0.382		0.809			1.207		
P值		0.709	0.537		0.668			0.547		

组别	年龄 例(%)		肿瘤分期 例(%)		高血压 例(%)		糖尿病 例(%)		吸烟 例(%)	
	< 60岁	≥ 60岁	I、II	III、IV	否	是	否	是	否	是
训练组	141(48.96)	147(51.04)	124(43.06)	164(56.94)	220(76.39)	68(23.61)	248(86.11)	40(13.89)	162(56.25)	126(43.75)
预测组	79(55.24)	64(44.76)	70(48.95)	73(51.05)	106(74.13)	37(25.87)	129(90.21)	14(9.79)	69(48.25)	74(51.75)
χ^2 值	1.270		1.114		0.157		1.115		2.147	
P值	0.260		0.291		0.692		0.291		0.143	

组别	饮酒 例(%)		口腔环境 例(%)		手术 例(%)		同步化疗 例(%)		Kamofsky 例(%)	
	否	是	否	是	否	是	否	是	否	是
训练组	147(51.04)	141(48.96)	128(44.44)	160(55.56)	177(61.46)	111(38.54)	233(80.9)	55(19.1)	162(56.25)	126(43.75)
预测组	65(45.45)	78(54.55)	64(44.76)	79(55.24)	84(58.74)	59(41.26)	125(87.41)	18(12.59)	79(55.24)	64(44.76)
χ^2 值	0.980		0.000		0.193		2.434		0.009	
P值	0.322		1.000		0.661		0.119		0.924	

组别	白细胞 例(%)		中性粒细胞计数 例(%)		红细胞计数 例(%)		血红蛋白 例(%)		白蛋白 例(%)	
	<4.0 × 10 ⁹ /L	≥4.0 × 10 ⁹ /L	<1.5 × 10 ⁹ /L	≥1.5 × 10 ⁹ /L	<4.3 × 10 ¹² /L	≥4.3 × 10 ¹² /L	<130 g/L	≥130 g/L	<35 g/L	≥35 g/L
训练组	31(10.76)	257(89.24)	13(4.51)	275(95.49)	114(39.58)	174(60.42)	136(47.22)	152(52.78)	41(14.24)	247(85.76)
预测组	24(16.78)	119(83.22)	10(6.99)	133(93.01)	61(42.66)	82(57.34)	70(48.95)	73(51.05)	20(13.99)	123(86.01)
χ^2 值	2.593		0.724		0.258		0.056		0.000	
P值	0.107		0.395		0.612		0.814		1.000	

归模型, XGBoost 模型的 AUC 为 0.835, 大于多因素 Logistic 回归模型的 0.830, 通过 PROC 包中的 roc.test 函数比较, 差异无统计学意义 ($Z = 0.136, P = 0.892$)

(见表 4)。多因素 Logistic 回归模型和 XGBoost 模型均有较好的预测准确性, 且预测效能相当。见图 3。

表 4 2 种模型的预测性能分析

模型	准确性/ %	95% CI		敏感性/ %	95% CI		特异性/ %	95% CI		阳性预 测值/%	阴性预 测值/%	AUC	95% CI	
		下限	上限		下限	上限		下限	上限				下限	上限
多因素 Logistic 回归模型	70.63	0.703	0.709	88.24	0.729	1.000	68.25	0.601	0.764	27.27	97.73	0.830	0.751	0.909
XGBoost 模型	71.33	0.711	0.723	82.35	0.642	1.000	69.84	0.627	0.786	26.92	96.70	0.835	0.741	0.929

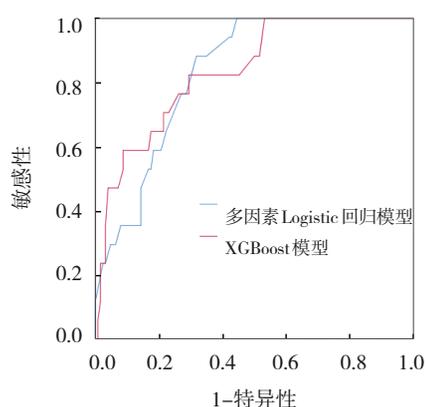


图 3 多因素 Logistic 回归模型和 XGBoost 模型的 ROC 曲线

3 讨论

放疗在有效遏制肿瘤的同时伴随多种并发症^[10-12], 口腔感染是舌癌放疗过程中常见的并发症^[13-14]。口腔癌放疗期间合并口腔感染, 随着感染时间延长, 癌细胞可能对放疗产生耐受, 将影响口腔癌患者放疗效果。如何能降低患者治疗期间不良反应, 降低口腔感染的发生率, 本文拟构建舌癌患者放疗期间口腔感染的预测模型, 评估舌癌患者口腔感染风险, 筛选出放疗期间口腔感染易感人群, 提早加强护理指导, 健康宣教, 采取预防措施, 降低口腔感染风险。本研究对影响舌癌患者放疗期间口腔感染的主要因素分别进行讨论。年龄: 多数人对口腔癌患者放疗期间发生口腔感染的危险因素分析显示, 年龄 > 60 岁是独立危险因素^[7-8]。本研究结论与之相同。考虑可能与老年患者机体免疫力低, 合并基础疾病风险高, 加之舌癌的局部侵袭损害, 使口腔癌患者放疗期间更容易出现口腔感染。肿瘤分期: 祁颖秋^[9]等研究显示, 头颈部肿瘤患者 T 分期越大, 患者

放疗期间发生口咽细菌感染风险越大。本研究同样显示口腔癌 T₃、T₄ 患者放疗期间发生口腔感染风险较 T₁、T₂ 患者高。考虑可能与肿瘤负荷较大有关。一般肿瘤负荷较大的患者, 患者靶区范围就会相应扩大, 放射线使患者口腔黏膜屏障受损的范围扩大。并且肿瘤负荷较大的患者更容易出现肿瘤坏死, 口腔清洁能力变差, 利于细菌真菌的滋生繁殖, 出现口腔感染。口腔环境: 吴春苏等^[15]研究提示, 鼻咽癌放疗期间, 差的口腔环境是影响患者口腔感染的独立危险因素。本研究同样显示口腔环境能力差的口腔癌患者放疗期间更容易出现口腔感染, 与大多数研结果相仿。放疗期间口腔黏膜损伤、口腔黏膜屏障受损、口腔环境改变、进食后未及时进行口腔清洁均易诱发口腔感染^[16-17]。手术、同步化疗: 王如等^[8]对影响口腔癌患者放疗期间发生口腔感染的因素进行分析, 同步化疗、未行手术为独立危险性因素。本研究在 XGBoost 模型中也提示手术、同步化疗是影响舌癌患者放疗期间口腔感染的重要因素。多因素 Logistic 回归模型提示手术是预防口腔感染的保护性因素, 手术患者发生口腔感染的风险较未手术患者低。未能行手术患者往往处于疾病晚期, 肿瘤本身易出现肿瘤坏死感染, 口腔生态环境受肿瘤影响破坏明显, 易引起患者口腔感染。考虑同步化疗降低患者免疫力, 并且化疗药物对患者口腔黏膜屏障有损害作用, 患者口腔黏膜抵抗能力下降, 易出现感染。血清白蛋白: 多数研究显示^[7-8, 18], 血清白蛋白 ≤ 40 g/L 的口腔癌患者放疗期间易出现口腔感染, 本研究结果同样显示血清白蛋白降低是发生口腔感染的独立危险性因素。舌癌影响患者进食, 加之肿瘤患者慢性消耗, 营养缺乏, 引起患者血清白蛋白下降, 免疫力降低, 诱发口腔感染^[19]。本研究还显示血红

蛋白为舌癌患者放疗期间发生口腔感染预后的独立保护性因素。有学者^[20]认为,组织血流灌注和氧合状态是抗感染免疫的重要组成部分,氧气可增强白细胞的有氧代谢,从而促进其杀灭病原体。血红蛋白的重要功能之一是为组织运输氧气,血红蛋白减少将影响组织供氧,机体的抗感染免疫下降。中性粒细胞计数:出现口腔感染的患者多伴有中性粒细胞计数下降^[18,18]。本研究在XGBoost模型中也提示中性粒细胞计数降低患者在放疗期间易出现口腔感染。中性粒细胞作为宿主抵抗入侵病原体的第一道防线,在激活后清除外来病原体,当中心粒细胞缺乏后,机体防御能力下降,出现口腔感染。

多因素 Logistic 回归模型和 XGBoost 模型比较分析:多因素 Logistic 模型是一种回归模型,适用于分类问题,可估计某个事件发生的可能性,也可分析某个问题的影响因素。依据多因素 Logistic 回归模型结果,通过 nomogram 绘制列线图,可更加直观地观察个体化预测结果^[21-22]。XGBoost 能实现准确的分类效果是通过弱分类器的迭代计算实现的。主要是降低偏差,也就是降低模型的误差。因此它是采用多个基学习器,每个基学习器都比较简单,避免过拟合,近年来广泛应用于医学领域^[23]。在预测组数据中两种模型的 ROC 曲线的 AUC 均>0.7,都表现出较好的预测准确性。进一步分析,XGBoost 模型的 AUC 略高于多因素 Logistic 回归模型,两者比较无差异,表明多因素 Logistic 回归模型和 XGBoost 模型预测效能相当,但 XGBoost 模型预测值略高于多因素 Logistic 回归模型。多因素 Logistic 回归模型与 XGBoost 模型的阳性预测值、准确性相近,说明 2 个模型正确分类的能力相当。XGBoost 模型的准确性和特异性均高于多因素 Logistic 模型,表明 XGBoost 模型实际预测能力较好,能很好地识别放疗期间发生口腔感染的舌癌患者。多因素 Logistic 模型的敏感性和阴性预测值均明显优于 XGBoost 模型。总体来说,XGBoost 模型和多因素 Logistic 模型均可预测舌癌患者放疗期间口腔感染的发生,2 个模型可取长补短。

XGBoost 模型变量重要性评分结果显示,口腔环境、手术、肿瘤分期、白蛋白、年龄、同步化疗、红细胞计数、血红蛋白、中性粒细胞计数为重要指标。多因素 Logistic 回归模型中独立危险因素为年龄、肿瘤分

期、口腔环境、手术、血清白蛋白,血红蛋白。以上指标对舌癌患者放疗期间发生口腔感染影响较大,尤其是口腔环境、手术、肿瘤分期、血清白蛋白等交叉指标。本研究可对存在以上因素的舌癌放疗患者发生口腔感染风险预警,对高危患者及早采取预防措施,降低口腔感染发生率,提高患者生活质量。

XGBoost 模型与多因素 Logistic 回归模型对舌癌患者放疗期间发生口腔感染的预测均有意义,两者预测能力相当。2 个模型纳入的预测因素均为临床常见指标,有利于在临床工作中开展筛查,积极做好预防口腔感染的工作,降低患者口腔感染的发生率。本研究的不足之处在于为单中心研究,数据量有限,今后尚需进行多中心,大样本的研究分析,以提高模型的预测精准性。

参 考 文 献 :

- [1] de BERARDINIS R, TAGLIABUE M, BELLONI P, et al. Tongue cancer treatment and oncological outcomes: the role of glossectomy classification[J]. *Surg Oncol*, 2022, 42: 101751.
- [2] MORIKAWA T, SHIBAHARA T, TAKANO M. Combination of fluorescence visualization and Iodine solution-guided surgery for local control of early tongue cancer[J]. *Int J Oral Maxillofac Surg*, 2023, 52(2): 161-167.
- [3] 郭建康,屠永奇,高振杰,等. 口腔鳞癌根治术后口腔感染病原菌及影响因素[J]. *中华医院感染学杂志*, 2022, 32(18): 2842-2846.
- [4] THAKKARP, BANKS JM, RAHATR, et al. Viruses of the oral cavity: prevalence, pathobiology and association with oral diseases[J]. *Rev Med Virol*, 2022, 32(4): e2311.
- [5] SROUSSI H Y, EPSTEIN J B, BENSADOUN R J, et al. Common oral complications of head and neck cancer radiation therapy: mucositis, infections, saliva change, fibrosis, sensory dysfunctions, dental caries, periodontal disease, and osteoradionecrosis[J]. *Cancer Med*, 2017, 6(12): 2918-2931.
- [6] CHITAPANARUX I, WONGSRITA S, SRIPAN P, et al. An underestimated pitfall of oral candidiasis in head and neck cancer patients undergoing radiotherapy: an observation study[J]. *BMC Oral Health*, 2021, 21(1): 353.
- [7] 郑莉萍. 口腔癌放疗患者发生口腔感染的危险因素[J]. *国际护理学杂志*, 2020, 39(19): 3503-3506.
- [8] 王如, 钱立庭, 高劲, 等. 口腔癌患者放疗后发生口腔感染的因素分析[J]. *现代仪器与医疗*, 2018, 24(6): 13-14.
- [9] 祁颖秋, 徐德静, 郑晓宇. 头颈部肿瘤患者放疗期间口咽细菌感染风险增加的护理研究[J]. *临床护理杂志*, 2015(5): 18-21.
- [10] TSAI C J, KUO Y H, WU H C, et al. Adjuvant radiotherapy significantly increases neck control and survival in early oral

- cancer patients with solitary nodal involvement: a national cancer registry database analysis[J]. *Cancers (Basel)*, 2021, 13(15): 3742.
- [11] MA B, HU Y, ZHU J D, et al. Research on the role of cellular autophagy in the sensitivity of human tongue cancer cells to radiotherapy and chemotherapy[J]. *J Stomatol Oral Maxillofac Surg*, 2023, 124(4): 101430.
- [12] SHANEI A, ABEDI I, SAADATMAND P, et al. Comparison of 3D conformal and intensity modulated radiotherapy in early stage oral tongue cancer: dosimetric and radiobiological evaluation[J]. *Int J Radiat Res*, 2020, 18(1): 33-42.
- [13] KAO Y S, HSU Y, HSU C Y. Radiotherapy increases the incidence of herpes zoster in oral cavity cancer patients - a national population-based cohort study[J]. *In Vivo*, 2021, 35(6): 3547-3553.
- [14] SINGH G K, CAPOOR M R, NAIR D, et al. Spectrum of fungal infection in head and neck cancer patients on chemoradiotherapy[J]. *J Egypt Natl Canc Inst*, 2017, 29(1): 33-37.
- [15] 吴春苏, 陈清华. 鼻咽癌放疗患者口腔感染的影响因素研究[J]. *癌症进展*, 2016, 14(5): 489-491.
- [16] SAUL-MCBETH J, DILLON J, LEE A, et al. Tissue damage in radiation-induced oral mucositis is mitigated by IL-17 receptor signaling[J]. *Front Immunol*, 2021, 12: 687627.
- [17] KAWASHITA Y, FUNAHARA M, YOSHIMATSU M, et al. A retrospective study of factors associated with the development of oral candidiasis in patients receiving radiotherapy for head and neck cancer: is topical steroid therapy a risk factor for oral candidiasis[J]. *Medicine (Baltimore)*, 2018, 97(44): e13073.
- [18] 罗利民, 贾勇士, 吴树强, 等. 口腔癌放疗患者口腔感染的危险因素调查[J]. *中华医院感染学杂志*, 2016, 26(7): 1609-1610.
- [19] BREWCZYŃSKI A, JABŁOŃSKA B, MROWIEC S, et al. Nutritional support in head and neck radiotherapy patients considering HPV status[J]. *Nutrients*, 2020, 13(1): 57.
- [20] SCHIETROMA M, CECILIA E M, DE SANTIS G, et al. Supplemental peri-operative oxygen and incision site infection after surgery for perforated peptic ulcer: a randomized, double-blind monocentric trial[J]. *Surg Infect (Larchmt)*, 2016, 17(1): 106-113.
- [21] SANTOS T R M, MELO J V, LEITE N C, et al. Usefulness of the vibration perception thresholds measurement as a diagnostic method for diabetic peripheral neuropathy: results from the Rio de Janeiro type 2 diabetes cohort study[J]. *J Diabetes Complications*, 2018, 32(8): 770-776.
- [22] JIANG S M, FANG J Y, YU T Y, et al. Novel model predicts diabetic nephropathy in type 2 diabetes[J]. *Am J Nephrol*, 2020, 51(2): 130-138.
- [23] CHEN T Q, GUESTRIN C. XGBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA, New York, NY, USA: Association for Computing Machinery, 2016: 785-794.

(张西倩 编辑)

本文引用格式: 张洋, 赵化荣, 刘攀, 等. 多因素 Logistic 回归模型和 XGBoost 模型对舌癌患者放疗期间发生口腔感染的预测价值[J]. *中国现代医学杂志*, 2023, 33(19): 66-73.

Cite this article as: ZHANG Y, ZHAO H R, LIU P, et al. Predictive value of logistic regression model and XGBoost model for the occurrence of oral infection during radiotherapy in patients with tongue cancer[J]. *China Journal of Modern Medicine*, 2023, 33(19): 66-73.